

Genome-wide association analysis of eosinophilic esophagitis provides insight into the tissue specificity of this allergic disease

Leah C Kottyan^{1–3,13}, Benjamin P Davis^{3,13}, Joseph D Sherrill³, Kan Liu³, Mark Rochman³, Kenneth Kaufman^{1,2}, Matthew T Weirauch^{1,4}, Samuel Vaughn¹, Sara Lazaro^{1,2}, Andrew M Rupert⁴, Mojtaba Kohram⁴, Emily M Stucke³, Katherine A Kemme³, Albert Magnusen^{1,2}, Hua He⁵, Phillip Dexheimer⁴, Mirna Chehade⁶, Robert A Wood⁷, Robbie D Pesek⁸, Brian P Vickery⁹, David M Fleischer¹⁰, Robert Lindblad¹¹, Hugh A Sampson⁶, Vincent A Mukkada¹², Phil E Putnam¹², J Pablo Abonia³, Lisa J Martin⁵, John B Harley^{1,2,14} & Marc E Rothenberg^{3,14}

Eosinophilic esophagitis (EoE) is a chronic inflammatory disorder associated with allergic hypersensitivity to food. We interrogated >1.5 million genetic variants in EoE cases of European ancestry and subsequently in a multi-site cohort with local and out-of-study control subjects. In addition to replicating association of the 5q22 locus (meta-analysis $P = 1.9 \times 10^{-16}$), we identified an association at 2p23 spanning *CAPN14* ($P = 2.5 \times 10^{-10}$). *CAPN14* was specifically expressed in the esophagus, was dynamically upregulated as a function of disease activity and genetic haplotype and after exposure of epithelial cells to interleukin (IL)-13, and was located in an epigenetic hotspot modified by IL-13. Genes neighboring the top 208 EoE-associated sequence variants were enriched for esophageal expression, and multiple loci for allergic sensitization were associated with EoE susceptibility ($4.8 \times 10^{-2} < P < 5.1 \times 10^{-11}$). We propose a model to explain the tissue-specific nature of EoE that involves the interplay of allergic sensitization with an EoE-specific, IL-13-inducible esophageal response involving *CAPN14*.

EoE is a chronic, food antigen-driven and tissue-specific inflammatory disease of the esophagus characterized by marked mucosal eosinophil accumulation that is often associated with fibrosis, stricture formation and impaired motility^{1–4}. The disease remits after removal of specific food types, reoccurs upon food reintroduction, is associated with marked dysregulation of the levels of esophageal transcripts rich in elements involved in allergic inflammation and can be induced

in mice by allergen exposure through IL-5- and IL-13-driven pathways^{5–8}. Consistent with an allergic etiology, EoE frequently co-occurs with allergic diseases, including asthma, eczema and food anaphylaxis^{2–4}. Why individuals with EoE develop a tissue-specific response remains an enigma, as the currently identified inflammatory pathways and genes in EoE overlap with those for other allergic diseases.

The only EoE genome-wide association study (GWAS) reported thus far identified a single susceptibility locus with significant association at 5q22, which harbors the *TSLP* gene encoding thymic stromal lymphopoietin⁹. A candidate gene analysis confirmed the association of EoE with *TSLP*, as well as with its receptor, *CRLF2* (encoding cytokine receptor-like factor 2)^{10,11}. Candidate gene studies have shown that *CCL26* (encoding eotaxin 3) and *FLG* (encoding filaggrin)^{12,13} are associated with EoE susceptibility. However, genetic variations in these genes and at 5q22 have also been associated with other atopic disorders^{14–16}, highlighting the need to elucidate how genetics might contribute to the tissue specificity of EoE.

Accordingly, we performed a GWAS of SNPs from >1.5 million genetic markers. This GWAS greatly expanded the number of EoE cases (from 351 to 736) and controls (from 3,105 to 9,246) over the previous study. Combined genetic association analysis identified 20 associated SNPs at 17 loci (threshold $P < 1 \times 10^{-7}$). Ten of these loci were identified in the analysis of our Cincinnati Children's Hospital Medical Center (CCHMC) cohort ($P < 1 \times 10^{-4}$) and in an analysis in which the external controls were excluded ($P < 5 \times 10^{-2}$) and were also independently confirmed in the US National Institutes of Health (NIH) Consortium of Food Allergy Research (CoFAR) cohort

¹Center for Autoimmune Genomics and Etiology, Department of Pediatrics, Division of Rheumatology, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio, USA. ²US Department of Veterans Affairs Medical Center, Cincinnati, Ohio, USA. ³Department of Pediatrics, Division of Allergy and Immunology, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio, USA. ⁴Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio, USA. ⁵Department of Pediatrics, Division of Human Genetics, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio, USA. ⁶Department of Pediatrics, Icahn School of Medicine at Mount Sinai, New York, New York, USA. ⁷Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ⁸Department of Pediatrics, University of Arkansas for Medical Sciences and Arkansas Children's Hospital, Little Rock, Arkansas, USA. ⁹Department of Pediatrics, University of North Carolina, Chapel Hill, North Carolina, USA. ¹⁰Department of Pediatrics, National Jewish Health, Denver, Colorado, USA. ¹¹The EMMES Corporation, Rockville, Maryland, USA. ¹²Department of Pediatrics, Division of Gastroenterology, Hepatology and Nutrition, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio, USA. ¹³These authors contributed equally to this work. ¹⁴These authors jointly directed this work. Correspondence should be addressed to M.E.R. (rothenberg@cchmc.org).

Received 13 March; accepted 18 June; published online 13 July 2014; doi:10.1038/ng.3033

LETTERS

Figure 1 Manhattan plot of the P values obtained from the genome-wide association analysis. (a) Data are from 736 subjects with EoE and 9,246 controls over 1,468,075 genetic variants with MAFs greater than 1% in the subjects with EoE. The $-\log_{10}$ value of each probability is shown as a function of genomic position on the autosomes. Genome-wide significance (red dashed line; $P \leq 5 \times 10^{-8}$) and suggestive significance (solid blue line; $P \leq 1 \times 10^{-7}$) are indicated. (b–e) Genetic association of variants at the 2p23 (b), 5q22 (c), 8p23 (d) and 15q13 (e) loci with EoE risk. P values ($-\log_{10}$) from the genetic association analysis of variants that were genotyped and imputed variants are plotted against the genomic position of each genotyped (blue) and imputed (red) SNP on the x axis on chromosomes 2, 5, 8 and 15. Genes in the region are shown above. Black lines indicate the recombination rates determined using subjects of European ancestry from the 1000 Genomes Project.

($P < 5 \times 10^{-2}$; Fig. 1, Table 1, Supplementary Fig. 1 and Supplementary Tables 1 and 2). This series of consistency checks was used to reduce the chance of identifying false positives due to the inclusion of an external control cohort.

Markers at 2p23, 5q22, 8p23 and 15q13 had associations that reached genome-wide significance ($P < 5 \times 10^{-8}$). The variants most highly associated with increased risk of EoE were found at 2p23 spanning the *CAPN14* gene (best SNP rs77569859, $P = 3.30 \times 10^{-10}$, odds ratio (OR) = 1.98) (Table 1). To identify commonly occurring variants (minor allele frequency (MAF) > 1%) on the risk haplotype that could be driving the genetic association, we imputed this region to a composite reference panel from the 1000 Genomes Project^{17,18}; no haplotype of continuous SNPs or haplotype constructed using the

most associated variants in the region was more highly associated with EoE risk than rs77569859 alone (best haplotype $P = 3.5 \times 10^{-8}$, OR = 1.6; Fig. 1b and data not shown). Variants at the other two newly discovered loci reaching genome-wide significance were located in the *XKR6* (encoding XK, Kell blood group complex subunit-related

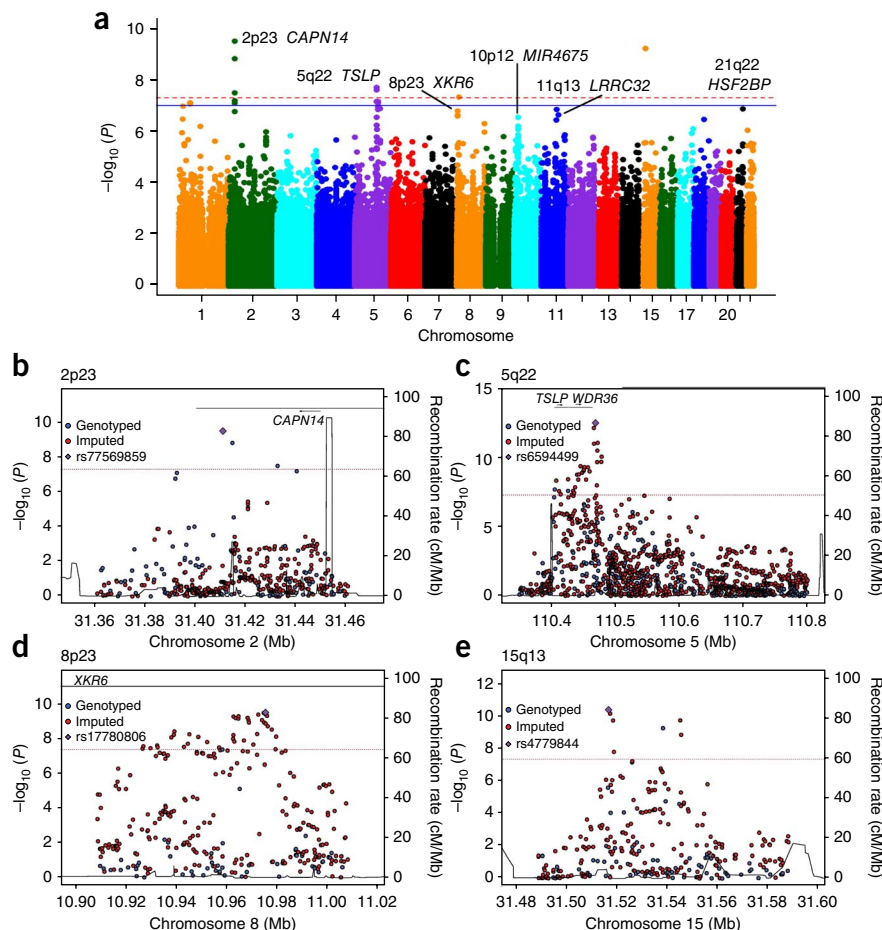


Table 1 Most highly associated EoE risk variants

Chr. band	SNP	Position (bp)	Minor allele	MAF cases	MAF controls	Major allele	P	OR	Nearest gene	Genomic context	P CCHMC	OR CCHMC	P CoFAR	OR CoFAR
1p13	rs2000260	108,673,405	G	0.37	0.43	A	6.56×10^{-7}	0.76	<i>SLC25A24</i>		1.10×10^{-4}	0.773	1.7×10^{-3}	0.725
2p23	rs77569859	31,411,287	G	0.09	0.05	A	3.30×10^{-10}	1.98	<i>CAPN14</i>	Intronic	1.41×10^{-7}	1.93	1.2×10^{-3}	2.06
5q22	rs3806933	110,406,742	A	0.37	0.44	G	2.00×10^{-8}	0.73	<i>TSLP</i>	Noncoding	1.32×10^{-6}	0.724	3.1×10^{-3}	0.740
5q23	rs2055376	116,181,428	A	0.04	0.02	C	7.12×10^{-8}	2.30	Near <i>SEMA6A</i>		1.89×10^{-6}	2.29	7.7×10^{-3}	2.40
8p23	rs2898261	10,958,539	A	0.35	0.42	C	4.84×10^{-8}	0.74	<i>XKR6</i>	Intronic	1.73×10^{-6}	0.724	6.2×10^{-3}	0.755
10p12	rs11819199	20,865,157	G	0.09	0.06	A	2.89×10^{-7}	1.62	<i>MIR4675</i>		1.64×10^{-5}	1.62	2.00×10^{-3}	1.70
11q13	rs2155219	76,299,194	C	0.41	0.49	A	3.65×10^{-7}	0.73	Between <i>C11orf30</i> and <i>LRRC32</i>		2.64×10^{-5}	0.743	5.8×10^{-3}	0.699
11q14	rs118086209	86,104,495	C	0.03	0.02	A	2.35×10^{-7}	2.19	<i>CCDC81</i>	Intronic	1.93×10^{-4}	1.99	6.61×10^{-5}	2.83
15q13	rs8041227	31,538,542	A	0.20	0.28	G	6.34×10^{-10}	0.66	Between <i>LOC283710</i> and <i>KLF13</i>		3.97×10^{-6}	0.693	3.38×10^{-5}	0.581
21q22	rs17004598	45,078,556	C	0.03	0.01	A	1.37×10^{-7}	2.57	<i>HSF2BP</i>	Intronic	4.48×10^{-6}	2.59	3.3×10^{-3}	2.75

Loci included in this table had significant or suggestive associations ($P < 5 \times 10^{-6}$) and were consistently associated in the CCHMC cohort ($P < 0.0005$), local analysis ($P < 0.05$) and CoFAR cohort ($P < 0.05$). Data for those loci that failed one of these consistency checks can be found in Supplementary Table 2. The most highly associated variant is shown for each locus. Position is given relative to Build 37 of the reference genome. The weighted z-score method implemented in METAL was used to combine the P values for the CCHMC and CoFAR cohorts. CCHMC, Cincinnati Children's Hospital Medical Center; CoFAR, US NIH Consortium of Food Allergy Research; chr. band, cytogenetic band; MAF, minor allele frequency across cases or controls; nearest gene, gene spanning (<1 Mb away from) the index SNP; OR, odds ratio for the minor allele; SNP, rsID of the variant. Bold font indicates loci with genome-wide significant association.

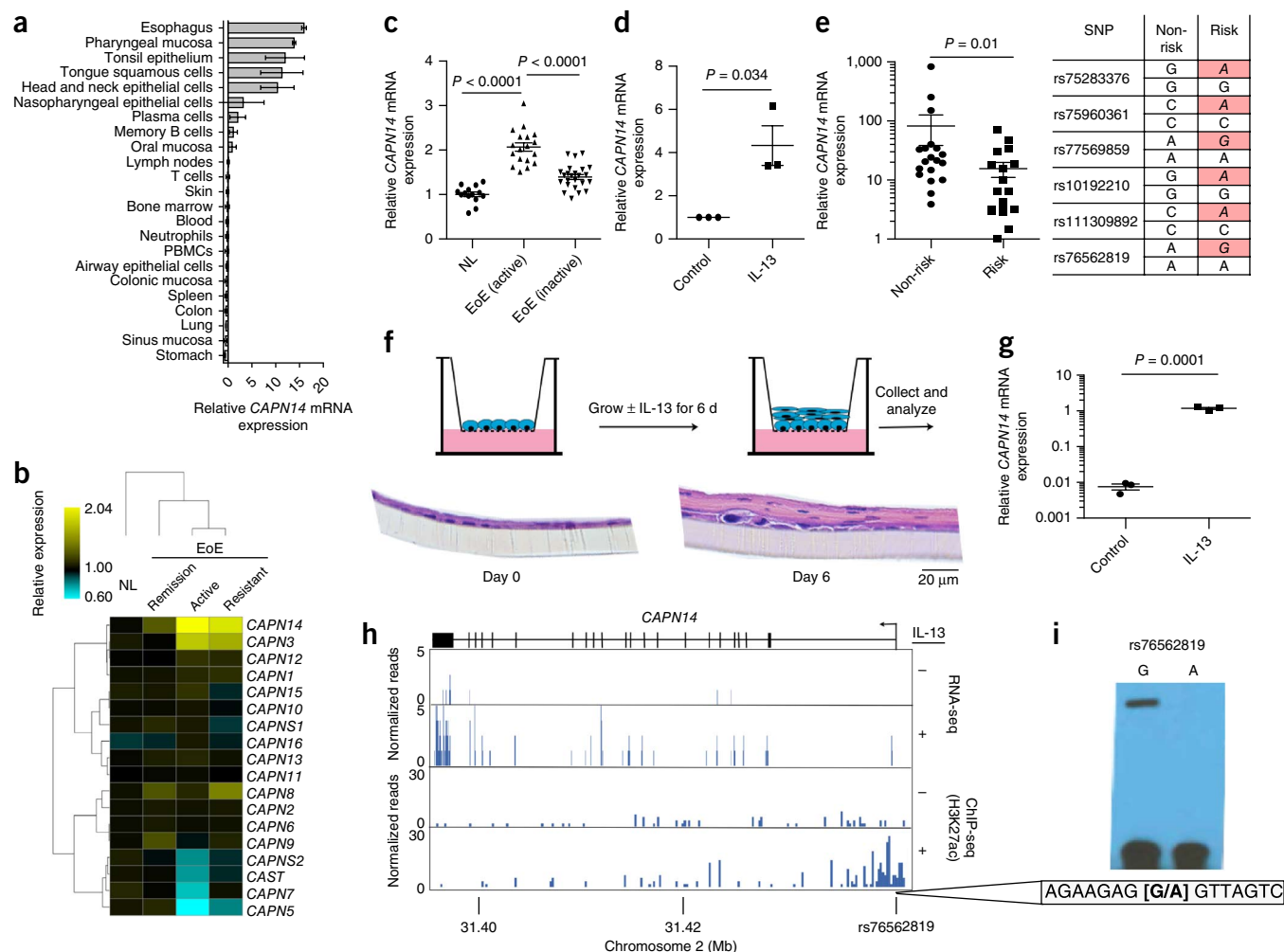


Figure 2 *CAPN14* is specifically expressed in esophageal epithelium and is dynamically upregulated as a function of disease activity and genetic haplotype and after exposure of epithelial cells to IL-13. (a) Barcode z-score relative microarray expression of *CAPN14* in various human tissue samples. Barcode analysis of 18,656 publicly available microarrays is displayed using <http://biogps.org/> (refs. 52,53). Data are representative from multiple cellular subtypes. Error bars represent the median absolute deviation. (For an exhaustive list of assessed tissues, see **Supplementary Fig. 2**.) (b) Microarray expression heat map of the calpain family in esophageal biopsies from normal controls (NL; $n = 14$), therapy-responsive EoE cases (EoE remission; $n = 18$), active EoE cases (EoE active; $n = 18$) and therapy-non-responsive EoE cases (EoE resistant; $n = 19$). (c,d) Microarray expression analysis of *CAPN14* expression from esophageal biopsies (NL, $n = 14$; EoE active, $n = 18$; EoE inactive, $n = 18$) (where n is the number of esophageal biopsies from different subjects) (c) and primary esophageal epithelial cells with or without IL-13 stimulation for 48 h ($n = 3$ biological replicates) (d). Error bars, s.e.m. (e) RT-PCR analysis of *CAPN14* expression relative to the expression of *HPRT1* (encoding hypoxanthine phosphoribosyltransferase 1) in biopsies from EoE cases with the non-risk haplotype ($n = 19$) or with at least one copy of the risk haplotype at the 2p23 locus ($n = 17$). The risk haplotype is defined as having the EoE risk allele (highlighted in red) at each of the six most highly associated variant locations. Error bars, s.e.m. (f) Schematic of the esophageal epithelial ALI Transwell culture system and hematoxylin and eosin staining demonstrating stratification. (g) RNA-seq expression analysis of *CAPN14* expression from ALI cultures with or without IL-13 stimulation for 6 d ($n = 3$ for each group). Error bars, s.e.m. (h) ChIP-seq analysis of the human esophageal cell line TE-7 shows increased levels of H3K27ac marks with IL-13 stimulation over the transcriptional start site of *CAPN14*, which correlate with an increase in transcriptional activity in RNA-seq analysis. One significantly associated SNP (rs76562819) is in this acetylation region. (i) EMSA was used to probe nuclear lysates from an esophageal epithelial cell line using oligonucleotides with the risk (G) or non-risk (A) allele of rs76562819.

family, member 6) gene (8p23) and in a gene desert (15q13). Very little is known about *XKR6*; however, public expression databases report its expression in the immune compartment (**Supplementary Fig. 2**). These four genome-wide susceptibility loci remained associated with EoE, and their effect sizes were not significantly influenced by correction for atopy (**Supplementary Table 3**).

Using an independently ascertained cohort that did not overlap with the first EoE GWAS⁹, we observed strong replication of disease association at 5q22 (rs6594499, Fisher's combined $P = 1.9 \times 10^{-16}$; **Supplementary Table 4**). After imputation of the region, the most

significant association with the development of EoE was downstream of *TSLP* and *WDR36* at rs1438672 (**Fig. 1c**), with 12 variants having association $P < 1 \times 10^{-2}$ (P values between 0.001 and 0.05) after adjusting for the most significant variant. After accounting for multiple testing, we cannot reject the null hypothesis that there is one independent genetic effect at this locus.

Variants at 1p13, 5q23, 10p12, 11q13, 11q14 and 21q22 demonstrated suggestive genetic association with EoE ($P < 1 \times 10^{-7}$) (**Fig. 1, Table 1 and Supplementary Figs. 3 and 4**). After establishing statistical associations between genetic variants at these loci and EoE, we

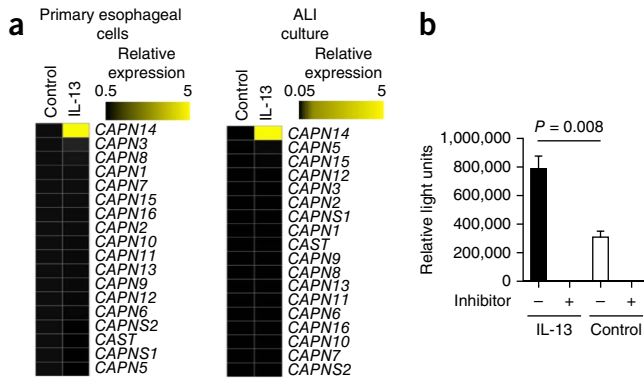


Figure 3 IL-13 induces *CAPN14* expression and calpain activity. (a) Heat map of microarray expression of the calpain family in primary esophageal epithelial cell culture (left) and in human esophageal progenitor cell line EPC2 ALI cultures (right). (b) Calpain activity assay in EPC2 cells with or without IL-13 stimulation for 48 h in the presence or absence of the calpain inhibitor PD150606. Error bars, s.e.m. (data are representative of three independent experiments). Light units are directly proportional to calpain enzyme activity.

performed fine-mapping studies starting with imputation of variants not captured in the combined GWAS data set (Fig. 1 and Supplementary Fig. 3). The 11q13 region has been associated with asthma^{19,20}, atopic dermatitis^{21–24}, inflammatory bowel disease²⁵, allergic rhinitis²⁶ and sensitization to grass²⁶. The EoE-associated variants at 11q13 were located between *C11orf30* (encoding chromosome 11 open reading frame 30) and *LRRC32* (encoding leucine-rich repeat-containing 32, also known as GARP). *LRRC32* has a role in latent transforming growth factor (TGF)- β surface expression²⁷, and *LRRC32* mRNA is highly expressed in activated forkhead box P3 (FOXP3)-positive T regulatory cells. Notably, TGF- β and FOXP3⁺ T regulatory cells have been implicated in EoE^{28–32}.

CAPN14 encodes calpain 14, a calcium-activated cysteine protease. A survey of 130 tissues showed that *CAPN14* was most highly expressed in the esophagus (Fig. 2a and Supplementary Fig. 5). When we assessed the expression of calpain family members in biopsies from subjects with and without EoE, we found distinct patterns for control subjects, those treated for EoE and those with uncontrolled (active) EoE. Notably, *CAPN14* was dynamically expressed as a function of disease activity (Fig. 2b). *CAPN14* showed the greatest upregulation in comparison to all members of the CAPN family (Fig. 2b), but the expression of 3 of the other 15 family members—*CAPN3*, *CAPN5* and *CAST* (encoding calpastatin)—was also dysregulated in EoE esophageal biopsies (Fig. 2b). *CAST* encodes a calpain inhibitor and was downregulated (29%; $P < 1 \times 10^{-4}$). We found a >2-fold increase in *CAPN14* levels in the esophageal biopsies of individuals with active EoE (Fig. 2c). Furthermore, IL-13 stimulation of primary esophageal epithelial cells and an esophageal epithelial cell line grown at the air-liquid interface (ALI) with IL-13 (Fig. 2f) resulted in a 4-fold and >100-fold increase in *CAPN14* levels, respectively (Fig. 2d,g). Individuals with the risk haplotype expressed 30% lower *CAPN14* mRNA levels than those without the risk haplotype ($P < 1 \times 10^{-2}$) (Fig. 2e).

To identify the genomic mechanisms that might underlie the association at 2p23 with EoE, we performed chromatin immunoprecipitation sequencing (ChIP-seq) on esophageal epithelial cells stimulated with IL-13. Of the six SNPs most highly associated at the *CAPN14* locus after imputation, two (rs76562819 and rs75960361) were located in putative regulatory regions, according to IL-13-induced acetylation of histone H3 at lysine 27 (H3K27ac) in ChIP-seq data for esophageal epithelial

cells, as well as publically available Encyclopedia of DNA Elements (ENCODE) functional genomics data (including histone marks, DNase I hypersensitivity data and ChIP-seq data)^{33–35} (Fig. 2h).

rs76562819 is located proximal to the 5' end of the *CAPN14* transcription start site (Fig. 2h), lies within a region of elevated mono-methylation of histone H3 at lysine 4 (H3K4me1) in multiple cell lines and intersects with open chromatin regions in 34 cell types, according to DNase I hypersensitivity site mapping data. We subsequently performed an electrophoretic mobility shift assay (EMSA) using a capture probe from this region and found that the risk allele (rs7462819) preferentially bound to a nuclear protein complex in comparison to the non-risk allele (Fig. 2i).

Notably, *CAPN14* was the only CAPN family member whose expression was upregulated, as measured by microarray and RNA sequencing (RNA-seq), in either primary epithelial cells from esophageal biopsies or organotypic cultures after stimulation with IL-13 (Fig. 3a), consistent with previous findings in other cell types^{36,37}. IL-13 increased calpain activity in esophageal epithelial cells stimulated with IL-13, and this activity was inhibited by a calpain-specific inhibitor (Fig. 3b).

We searched for proximal (*cis*-acting) expression quantitative trait loci (eQTLs) using gene expression data obtained from six cell types or tissues (white blood cells, lymphoblastoid cell lines, whole blood, adipose tissue, B cells and monocytes). At two of the ten most highly associated EoE risk loci, the sentinel SNP was associated ($P < 1 \times 10^{-3}$) with the expression of one or more nearby genes. Specifically, the most highly associated variants at 1p13 and 8p23 were strongly associated with the expression of the nearest genes, *SLC25A24* (encoding solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 24; $P = 1.25 \times 10^{-9}$) and *XKR6* ($P = 1.02 \times 10^{-7}$) (Supplementary Table 5). Using our own expression databases, we have previously reported that *TSLP* expression is upregulated in biopsies from individuals with EoE in an allele-dependent manner and in esophageal epithelial cell lines treated with polyinosinic:polycytidylic acid^{9,10}.

We measured the expression in RNA-seq data of all genes within 25 kb of variants with combined association $P < 1 \times 10^{-4}$. Of these 208 genes, 48% were expressed in the esophagus at appreciable levels³⁸; this proportion represents an enrichment relative to the expression of the whole genome in the esophagus ($P < 1 \times 10^{-4}$, permutation testing). Furthermore, differential expression of these genes was sufficient to segregate EoE cases from controls, and 8% of the genes near the most highly associated EoE risk loci were differentially expressed (>2-fold average change, Bonferroni-adjusted $P < 5 \times 10^{-2}$) in EoE case

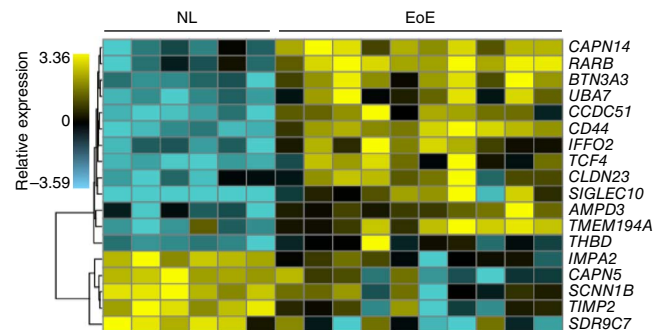


Figure 4 Genes at EoE risk loci with differential expression in biopsies from EoE cases relative to controls. Genes within 25 kb of the 768 genetic variants associated with EoE (combined $P < 1 \times 10^{-4}$) were used in this analysis of RNA-seq data. The expression of 208 genes was assessed. Normalized fold change in expression is shown for all genes with at least twofold average change and corrected $P < 0.05$. NL, normal controls.

and control biopsies (Fig. 4 and Supplementary Fig. 6). Therefore, these data demonstrate a concentration of EoE susceptibility loci in the neighborhood of genes that are expressed and/or dysregulated in the esophagus of diseased individuals, suggesting an esophageal functional role for the implicated genes. We assessed the GWAS associations for differential H3K27ac marks in epithelial cells after IL-13 stimulation. Notably, *CAPN14* was one of two genes with these IL-13-responsive epigenetic characteristics (Supplementary Fig. 7). This analysis further underscores the potential centrality of *CAPN14* in the etiology of EoE. These data are consistent with the mechanistic model in which *CAPN14* is dynamically expressed in the esophagus in response to inflammatory stimuli, a regulatory mechanism that is disrupted by the decreased expression associated with the risk haplotype (Supplementary Fig. 8).

The previous genome-wide study of EoE did not assess the most significant variant in *CAPN14* but did identify suggestive association ($P < 1 \times 10^{-4}$) from variants in the region⁹. With additional EoE cases and controls, the current study was better powered to identify statistically significant association of genetic variants with the development of EoE. A limitation of this study is the lack of a replication stage for those loci that were most highly associated with EoE in the combined analysis of the two independent cohorts that were assessed. As such, apart from the associations at *TSLP-WDR36* and *CAPN14*, the highly associated loci remain suggestive until confirmed in subsequent studies.

Two recent GWAS reported 22 loci associated with allergic sensitization^{39,40}. Notably, we found association with EoE at 9 of these 22 loci (Supplementary Table 6), underscoring the key role of atopy in EoE, and 8 of these SNPs were associated with comparable effects on disease risk. The atopic sensitization loci with the greatest association with EoE were located in *CLEC16A* (encoding C-type lectin domain family 16, member A), *LRRC32* (encoding leucine-rich repeat-containing 32), *LPP* (encoding LIM domain-containing preferred translocation partner in lipoma) and *TSLP-WDR36* (Fig. 1 and Supplementary Fig. 9). However, of the 10 loci that were associated with EoE in our study, only 2 overlapped with the 22 allergic sensitization loci, highlighting the possibility that non-atopy-related processes might be operational in EoE.

In conclusion, herein we have increased the number of putative susceptibility loci for EoE from one to four and present compelling evidence for six other loci. Notably, our data substantiate a mechanism to elucidate the tissue-specific manifestations of this prototypic allergic disease. In particular, we provide additional evidence of shared genetic and molecular pathways in general atopy risk factors (for example, *TSLP-WDR36*, *LRRC32*, *IL33* and *LPP*) and EoE disease-specific elements, most notably genetic risk factors present at 2p23 where *CAPN14* is located. Consistent with this model, the *CAPN14* gene is located in a baseline epigenetic hotspot that is modified by IL-13, and *CAPN14* is specifically expressed in esophageal epithelium and is dynamically upregulated as a function of disease activity and genetic haplotype and after exposure of epithelial cells to IL-13. We further note that mutations in *CAPN3*, whose gene product is specifically expressed in skeletal muscle, have been associated with susceptibility to another tissue-specific eosinophilic disorder (eosinophilic myositis)^{41–46}. Although we do not yet know the molecular steps that link calpainopathy to eosinophilic responses, it is notable that *CAPN14* belongs to the classical calpain subfamily that comprises one of the major proteolytic systems that mediate protein cleavage (in addition to the proteasome, lysosome and caspase systems)⁴⁷. Classical calpains are calcium-regulatory proteases, and their substrates include structural proteins, signaling molecules, transcription factors^{48,49} and the inflammatory mediators that are germane for allergic

responses^{50,51}, such as STAT6 (signal transducer and activator of transcription 6) and IL-33, the latter of which shows some association with EoE (Supplementary Fig. 7 and Supplementary Table 2). On the basis of the collective data, we propose a model that links the interplay of allergic sensitization with an EoE-specific, IL-13-inducible esophageal response involving *CAPN14*.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. The genotyping data from this study have been submitted to the database of Genotypes and Phenotypes (dbGaP) under accession [phs000494.v1.p1](#). The expression data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO) database under accession [GSE58640](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We kindly thank the participants and their families. This study would not be possible without the CoFAR site investigators: S.H. Sicherer, S.M. Jones, T.T. Perry, A. Scurlock, E. Kim, E. Dellon, A.W. Burks, G.T. Furuta, D.Y.M. Leung and D. Stablein. Other people who contributed to this study include J. Gau, J. Fronda, D. Ananos, K. Mudd, T. Fitzgerald, P. Chandler, E. O'Brien, S. Carlisle, A. Hiegel, P. Steele, J. Gebhart, D. Hamilton, W. Moore, J. Newton, S. Leung, D. Brown, M. Plaut and J. Poyser.

Funding for this project was provided by the US NIH, National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Heart, Lung, and Blood Institute (NHLBI) and National Human Genome Research Institute (NHGRI) (U19 AI066738, U01 HG006828, U01 HG006828-S1, U01 HG006828-S2, U01 AI066560, R37 AI024717, P01 AI083194, T32 HL7752-19, K23 AI099083 and P01 AR049084). The CoFAR arm of the study was supported by US NIH grant U19 AI066738 from NIAID and NIDDK. The project was also supported by several grants from the National Center for Research Resources (NCRR), a component of the US NIH: UL1 TR001082 (National Jewish), UL1 TR-000067 (Mount Sinai), UL1 TR-000039 (Arkansas), UL1 TR-000083 (University of North Carolina) and UL1 TR-000424 (Johns Hopkins). Support was also received from the US Department of Veteran Affairs (IMMA 9) and the US Department of Defense (PR094002). This research was supported in part by the Cincinnati Children's Research Foundation and its Cincinnati Genomic Control Cohort. Other support was derived from the Campaign Urging Research for Eosinophilic Diseases (CURED), the Buckeye Foundation, the Food Allergy Research Education (FARE) Foundation and the Foundation of the American College of Allergy, Asthma and Immunology.

The paper's contents are solely the responsibility of the authors and do not necessarily represent the official view of the NCRR or NIH. We also thank S. Hottinger for editorial assistance.

AUTHOR CONTRIBUTIONS

M.E.R. and J.B.H. conceived the project and supervised the study. L.C.K. and M.E.R. wrote the manuscript. L.C.K. led and performed the genetic analysis. H.H. and L.J.M. contributed statistical consultation and support. B.P.D., J.D.S. and M.R. performed the experiments with assistance from E.M.S. and K.A.K., K.L., K.K., M.T.W., A.M.R., M.K., P.D. and S.V. provided analytical and bioinformatics support. S.L. genotyped samples with assistance from A.M., M.C., R.A.W., R.D.P., B.P.V., D.M.F., R.L. and H.A.S. provided samples through the CoFAR consortium. V.A.M., P.E.P., J.P.A. and M.E.R. provided samples for the CCHMC cohort. All of the authors discussed the results and commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the [online version of the paper](#).

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Liacouras, C.A. *et al.* Eosinophilic esophagitis: updated consensus recommendations for children and adults. *J. Allergy Clin. Immunol.* **128**, 3–20 (2011).

2. Rothenberg, M.E. Biology and treatment of eosinophilic esophagitis. *Gastroenterology* **137**, 1238–1249 (2009).
3. Collins, M.H. *et al.* Clinical, pathologic, and molecular characterization of familial eosinophilic esophagitis compared with sporadic cases. *Clin. Gastroenterol. Hepatol.* **6**, 621–629 (2008).
4. Putnam, P.E. & Rothenberg, M.E. Eosinophilic esophagitis: concepts, controversies, and evidence. *Curr. Gastroenterol. Rep.* **11**, 220–225 (2009).
5. Abonia, J.P. & Rothenberg, M.E. Eosinophilic esophagitis: rapidly advancing insights. *Annu. Rev. Med.* **63**, 421–434 (2012).
6. Davis, B.P. & Rothenberg, M.E. Emerging concepts of dietary therapy for pediatric and adult eosinophilic esophagitis. *Expert Rev. Clin. Immunol.* **9**, 285–287 (2013).
7. Henderson, C.J. *et al.* Comparative dietary therapy effectiveness in remission of pediatric eosinophilic esophagitis. *J. Allergy Clin. Immunol.* **129**, 1570–1578 (2012).
8. Rothenberg, M.E. *et al.* Working with the US Food and Drug Administration: progress and timelines in understanding and treating patients with eosinophilic esophagitis. *J. Allergy Clin. Immunol.* **130**, 617–619 (2012).
9. Rothenberg, M.E. *et al.* Common variants at 5q22 associate with pediatric eosinophilic esophagitis. *Nat. Genet.* **42**, 289–291 (2010).
10. Sherrill, J.D. *et al.* Variants of thymic stromal lymphopoietin and its receptor associate with eosinophilic esophagitis. *J. Allergy Clin. Immunol.* **126**, 160–165 (2010).
11. Zuo, L. *et al.* IL-13 induces esophageal remodeling and gene expression by an eosinophil-independent, IL-13R α 2-inhibited pathway. *J. Immunol.* **185**, 660–669 (2010).
12. Blanchard, C. *et al.* Coordinate interaction between IL-13 and epithelial differentiation cluster genes in eosinophilic esophagitis. *J. Immunol.* **184**, 4033–4041 (2010).
13. Blanchard, C. *et al.* Eotaxin-3 and a uniquely conserved gene-expression profile in eosinophilic esophagitis. *J. Clin. Invest.* **116**, 536–547 (2006).
14. McAleer, M.A. & Irvine, A.D. The multifunctional role of filaggrin in allergic skin disease. *J. Allergy Clin. Immunol.* **131**, 280–291 (2013).
15. Fallon, P.G. *et al.* A homozygous frameshift mutation in the mouse *Flg* gene facilitates enhanced percutaneous allergen priming. *Nat. Genet.* **41**, 602–608 (2009).
16. Shin, H.D. *et al.* Association of Eotaxin gene family with asthma and serum total IgE. *Hum. Mol. Genet.* **12**, 1279–1285 (2003).
17. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
18. Altshuler, D.M. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
19. Ferreira, M.A. *et al.* Identification of *IL6R* and chromosome 11q13.5 as risk loci for asthma. *Lancet* **378**, 1006–1014 (2011).
20. Tang, H.Y. *et al.* Association analysis of single nucleotide polymorphisms at five loci: comparison between atopic dermatitis and asthma in the Chinese Han population. *PLoS ONE* **7**, e35334 (2012).
21. Esparza-Gordillo, J. *et al.* A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat. Genet.* **41**, 596–601 (2009).
22. Greisenegger, E.K., Zimprich, F., Zimprich, A., Gleiss, A. & Kopp, T. Association of the chromosome 11q13.5 variant with atopic dermatitis in Austrian patients. *Eur. J. Dermatol.* **23**, 142–145 (2013).
23. Hirota, T. *et al.* Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. *Nat. Genet.* **44**, 1222–1226 (2012).
24. Paternoster, L. *et al.* Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. *Nat. Genet.* **44**, 187–192 (2012).
25. Barrett, J.C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **40**, 955–962 (2008).
26. Ramasamy, A. *et al.* A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization and their interaction with birth order. *J. Allergy Clin. Immunol.* **128**, 996–1005 (2011).
27. Tran, D.Q. *et al.* GARP (LRRC32) is essential for the surface expression of latent TGF- β on platelets and activated FOXP3⁺ regulatory T cells. *Proc. Natl. Acad. Sci. USA* **106**, 13445–13450 (2009).
28. Fuentesbella, J. *et al.* Increased number of regulatory T cells in children with eosinophilic esophagitis. *J. Pediatr. Gastroenterol. Nutr.* **51**, 283–289 (2010).
29. Stuck, M.C., Straumann, A. & Simon, H.U. Relative lack of T regulatory cells in adult eosinophilic esophagitis—no normalization after corticosteroid therapy. *Allergy* **66**, 705–707 (2011).
30. Tantibhaedhyangkul, U., Tatevian, N., Gilger, M.A., Major, A.M. & Davis, C.M. Increased esophageal regulatory T cells and eosinophil characteristics in children with eosinophilic esophagitis and gastroesophageal reflux disease. *Ann. Clin. Lab. Sci.* **39**, 99–107 (2009).
31. Zhu, X., Wang, M., Crump, C.H. & Mishra, A. An imbalance of esophageal effector and regulatory T cell subsets in experimental eosinophilic esophagitis in mice. *Am. J. Physiol. Gastrointest. Liver Physiol.* **297**, G550–G558 (2009).
32. Frischmeyer-Guerrero, P.A. *et al.* TGF β receptor mutations impose a strong predisposition for human allergic disease. *Sci. Transl. Med.* **5**, 195ra94 (2013).
33. Maher, B. ENCODE: the human encyclopaedia. *Nature* **489**, 46–48 (2012).
34. Ecker, J.R. *et al.* Genomics: ENCODE explained. *Nature* **489**, 52–55 (2012).
35. Skipper, M., Dhand, R. & Campbell, P. Presenting ENCODE. *Nature* **489**, 45 (2012).
36. Ueta, M., Mizushima, K., Yokoi, N., Naito, Y. & Kinoshita, S. Expression of the interleukin-4 receptor α in human conjunctival epithelial cells. *Br. J. Ophthalmol.* **94**, 1239–1243 (2010).
37. Ueta, M., Sotozono, C. & Kinoshita, S. Expression of interleukin-4 receptor α in human corneal epithelial cells. *Jpn. J. Ophthalmol.* **55**, 405–410 (2011).
38. Sherrill, J.D. *et al.* Analysis and expansion of the eosinophilic esophagitis transcriptome by RNA sequencing. *Genes Immun.* doi:10.1038/gene.2014.27 (12 June 2014).
39. Bønnelykke, K. *et al.* Meta-analysis of genome-wide association studies identifies ten loci influencing allergic sensitization. *Nat. Genet.* **45**, 902–906 (2013).
40. Hinds, D.A. *et al.* A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. *Nat. Genet.* **45**, 907–911 (2013).
41. Krahn, M. *et al.* Eosinophilic infiltration related to *CAPN3* mutations: a pathophysiological component of primary calpainopathy? *Clin. Genet.* **80**, 398–402 (2011).
42. Krahn, M. *et al.* *CAPN3* mutations in patients with idiopathic eosinophilic myositis. *Ann. Neurol.* **59**, 905–911 (2006).
43. Bartoli, M. *et al.* Validation of comparative genomic hybridization arrays for the detection of genomic rearrangements of the calpain-3 and dysferlin genes. *Clin. Genet.* **81**, 99–101 (2012).
44. Brown, R.H. Jr. & Amato, A. Calpainopathy and eosinophilic myositis. *Ann. Neurol.* **59**, 875–877 (2006).
45. Amato, A.A. Adults with eosinophilic myositis and calpain-3 mutations. *Neurology* **70**, 730–731 (2008).
46. Oflazer, P.S., Gundesli, H., Zorludemir, S., Sabuncu, T. & Dincer, P. Eosinophilic myositis in calpainopathy: could immunosuppression of the eosinophilic myositis alter the early natural course of the dystrophic disease? *Neuromuscul. Disord.* **19**, 261–263 (2009).
47. Sorimachi, H., Hata, S. & Ono, Y. Calpain chronicle—an enzyme family under multidisciplinary characterization. *Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci.* **87**, 287–327 (2011).
48. Arandis, T. *et al.* Differential functions of calpain 1 during epithelial cell death and adipocyte differentiation in mammary gland involution. *Biochem. J.* **459**, 355–368 (2014).
49. Wang, Y. & Zhang, Y. Regulation of TET protein stability by calpains. *Cell Rep.* **6**, 278–284 (2014).
50. Meephansan, J., Tsuda, H., Komine, M., Tominaga, S. & Ohtsuki, M. Regulation of IL-33 expression by IFN- γ and tumor necrosis factor- α in normal human epidermal keratinocytes. *J. Invest. Dermatol.* **132**, 2593–2600 (2012).
51. Hayakawa, M. *et al.* Mature interleukin-33 is produced by calpain-mediated cleavage *in vivo*. *Biochem. Biophys. Res. Commun.* **387**, 218–222 (2009).
52. Wu, C. *et al.* BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* **10**, R130 (2009).
53. McCall, M.N., Uppal, K., Jaffee, H.A., Zilliox, M.J. & Irizarry, R.A. The Gene Expression Barcode: leveraging public data repositories to begin cataloging the human and murine transcriptomes. *Nucleic Acids Res.* **39**, D1011–D1015 (2011).

ONLINE METHODS

Genotyping. Genotyping was performed on the Illumina OMNI-5 and OMNI-2.5 genotyping arrays using Infinium2 chemistry. Genotypes were called using the Gentrain2 algorithm in Illumina Genome Studio software. A summary of the genotyping and analytical pipeline is given in **Supplementary Figure 10**.

Subjects. The study was approved by the institutional review boards at CCHMC and all participating sites that were part of the US NIH CoFAR EoE cohort (Mount Sinai Medical Center, University of North Carolina, Johns Hopkins University, University of Colorado Health Center/National Jewish Research Center and Arkansas Children's Hospital). Parental informed consent was obtained from all participants in this study for the purpose of DNA collection and genotyping. Cases were confirmed by a physician to fulfill the diagnostic criteria for EoE. EoE is defined by a peak eosinophil count of ≥ 15 eosinophils/high-power field in esophageal biopsy sections; 30% of CCHMC and 51% of CoFAR cases had proton pump inhibitor (PPI) therapy before diagnostic endoscopy. Control subjects (without EoE) included the self-reported Caucasian subjects in the Cincinnati Genomic Control Cohort CCHMC ($n = 831$; age range of 2–18 years)⁵⁴ and an external control cohort (without EoE) acquired from a database of Genotypes and Phenotypes (dbGaP) University of Michigan study ($n = 8,580$). In the CCHMC and CoFAR cohorts, 73% and 62% of EoE cases were male, respectively, and EoE cases had an age range of 2–52 years. The external control cohort was collected through an aging and retirement study; subjects in this cohort were older than the cases at the time of DNA collection. The subjects in the external control cohort were randomly assigned to the CCHMC or CoFAR analysis with the goal of an equivalent case/control ratio in each cohort.

Population stratification. Genome-wide data were used to infer the top six principal components of genetic variation and correct for possible population stratification using EIGENSOFT. All local cases and controls were self-identified as having European ancestry, and principal-component analysis was used to exclude all subjects ($n = 271$) who segregated more than 4 s.d. outside of the mean of the first 4 principal components. Specifically, we removed 34 genetic outliers from 736 EoE cases, 13 outliers from 235 CoFAR cases, 71 outliers from 831 CCHMC controls and 166 outliers from 8,652 dbGaP University of Michigan controls. The resulting genomic inflation factor was 1.001 for a set of ancestry-informative markers not including the associated loci. The genome-wide genomic inflation factors seen in our study for the CCHMC and CoFAR cohorts were 1.04 and 1.05, respectively.

Genotyping quality control. To confirm accurate genotyping and sample identification, we performed an identity-by-state analysis on a subset of 10,732 SNPs that were genotyped on both the Illumina OMNI-5 array and the previous GWAS platform (Illumina 550)⁹ and found 99.998% concordance for the 134 samples that were genotyped on both platforms. Genome-wide genotyping data were deposited into dbGaP ([phs000494.v1.p1](https://www.ncbi.nlm.nih.gov/bioproject/1000000000)). Of the original 4,301,332 markers on the OMNI-5 array, 2,512,766 were derived from rare variants ($MAF < 1\%$). Notably, 87,539 markers had a suboptimal call rate (for example, $< 96\%$), and 44,246 were not in Hardy-Weinberg equilibrium in the controls ($P < 1 \times 10^{-4}$). We assessed the 10,000 most significant variants in Genome Studio projects that included the locally genotyped cases and controls and the external controls from dbGaP; upon visual inspection, 124 of these SNPs did not cluster properly and were removed from further analysis. rs8041227 reached genome-wide significance; however, no other SNPs around this variant in the data set were strongly associated. The Genome Studio cluster plot of this particular SNP is given in **Supplementary Figure 11**. External controls were genotyped on the OMNI-2.5 array, and overlapping variants that passed the quality measures described above in both data sets were analyzed. We controlled for the presence of potential batch effects in the analysis of the CoFAR samples by removing SNPs that exhibited outlier fluorescence associated with deviation between plates ($P < 1 \times 10^{-4}$), per the manufacturer's recommendation. The final genotyping rate for all SNPs was 99.7%. After applying the above filters, we used genotypes from 1,468,075 autosomal SNPs in a CCHMC cohort of 435 cases and 716 well-phenotyped local controls from CCHMC and 5,776 controls from the dbGaP study and a CoFAR cohort of 222 cases from the US NIH CoFAR EoE cohort and 2,804 controls from the University of Michigan Health and Retirement System (obtained from dbGaP) (**Supplementary Table 1**).

Association and linkage disequilibrium analysis. After removing genetic outliers, we performed logistic regression to calculate P values and OR estimates for each SNP using PLINK⁵⁵ with sex as a covariate. For some analyses, atopy and the most significant SNP in the region were also used as covariates. For analyses that used atopy as a covariate, only subjects with known atopic status (EoE cases and local Cincinnati controls) were included. For cases, atopy was defined by a physician-documented history of positive skin-prick test results, allergic rhinitis, allergic dermatitis/eczema, asthma or food allergy. The prevalence of atopy for the local CCHMC and CoFAR EoE cohorts was 96.2% and 91.3%, respectively. For the Cincinnati Genomic Control Cohort, atopy was defined by a parent-reported history of allergic rhinitis, eczema, asthma or food allergy; the atopy prevalence of this cohort was 28.6%. LocusZoom⁵⁶ and R were used to map the associated loci in the context of chromosomal recombination and nearby genes.

Imputation to the 1000 Genomes Project reference panel. To detect associated variants that were not directly genotyped, we imputed highly associated regions with IMPUTE2 and used a composite imputation reference panel of integrated haplotypes from the 1000 Genomes Project sequence data freezes from March 2012 produced using the SHAPEIT2 program^{17,18}. Imputed genotypes were required to meet or exceed a probability threshold of 0.9, an information measure of > 0.4 and the quality control criteria described for the genotyped markers.

Expression analysis. Quantitative RT-PCR analysis was performed on cDNA synthesized from distal esophageal biopsy RNA. EoE biopsies showed active disease pathology at the time they were taken, and all cases reported no glucocorticoid treatment at the time of biopsy, except for the analysis of gene expression as a function of disease activity in which remission was defined after glucocorticoid therapy. Statistical testing for mRNA expression normalized to the levels of *GAPDH* (encoding glyceraldehyde 3-phosphate dehydrogenase) or *HPRT1* (encoding hypoxanthine phosphoribosyltransferase 1) was performed by Mann-Whitney U test using GraphPad Prism software.

Expression microarray analysis. For each case, one distal esophageal mucosal biopsy sample was immersed in RNeasy RNA stabilization reagent (Qiagen) and stored at 4 °C for less than 15 d. Total RNA was extracted using the RNeasy Mini kit (Qiagen) according to the manufacturer's recommendations. Hybridization to DNA microarrays was performed by the Microarray Core at CCHMC, as previously reported^{13,57}. The genome-wide human Affymetrix U133 Plus 2.0 GeneChip was used, and the relative expression of gene transcript levels were determined using algorithms in the Microarray Analysis Suite and GeneSpring software (Silicon Genetics).

RNA sequencing. Esophageal biopsy RNA was isolated from EoE cases with active disease and from unaffected controls as previously described^{13,38}. RNA-seq acquiring 10 million mappable 100-bp reads from paired-end libraries was performed at the Genetic Variation and Gene Discovery Core Facility at CCHMC. Data were aligned to the GrCh37 build of the human genome using Ensembl⁵⁸ annotations as a guide for TopHat⁵⁹. Expression analysis was performed using Cufflinks⁶⁰. Data were visualized using the Integrative Genomics Viewer (Broad Institute). Enrichment analysis was performed using permutation testing that randomly selected SNPs from the chip, identified the nearest gene and assessed the number of genes expressed in the esophagus. Only 10 permutations out of 100,000 total permutations had more than or equal to 99 genes expressed in the esophagus.

Electrophoretic mobility shift assays. Pairs of single-stranded 5' IRDye infrared dye-labeled and unlabeled oligonucleotides (obtained from IDT, Inc.) were annealed to generate double-stranded probes. We incubated 25–50 fmole of labeled probe with 8 or 10 μ g of nuclear extract prepared from the human esophageal cell line TE-7 (tested for mycoplasma infection), 6 μ g of poly(deoxyinosinic-doxycytidylic) acid and 1 μ l of salmon sperm provided along with the buffers and protocols supplied with the Odyssey Infrared EMSA kit (LI-COR Biosciences). Binding reactions were analyzed using electrophoresis on 6% Tris-borate-EDTA polyacrylamide gels and detected by an infrared

fluorescent procedure using the Odyssey Infrared Imaging System (LI-COR Biosciences).

Organotypic cultures. For the ALI culture system, the esophageal epithelial cell line (the hTERT-immortalized EPC2 line from A. Rustgi (University of Pennsylvania)) was grown to confluence on polyester-permeable supports with a pore size of 0.4 μm (Corning) in keratinocyte serum-free medium (K-SFM) (Life Technologies) supplemented with 1.8 mM calcium. Epithelial cell differentiation was then induced by removing the culture medium from the inner chamber of the permeable support and maintaining the esophageal epithelial cells for 5–7 d at the ALI in the presence or absence of IL-13 (100 ng/ml).

H3K27ac analysis. We fixed 10–20 million TE-7 cells (a kind gift from P. Hainault, International Agency for Research on Cancer) with 0.8% formaldehyde by adding 1 ml of 10 $\times$ fixation buffer (50 mM HEPES-KOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 8% formaldehyde) to 9 ml of growth medium for 8 min at room temperature with shaking. The reaction was stopped by adding glycine to a final concentration of 125 mM for an additional 5 min. Nuclei were prepared with the truChIP High Cell Chromatin Shearing Kit with SDS Shearing Buffer (Covaris) according to the manufacturer's recommendations. Sonication was performed using a Covaris S220 Focused ultrasonicator at 175 pip, 10% output and 200 bursts for 8 min. Efficient DNA fragmentation was verified by agarose gel electrophoresis. ChIP was performed with 2 μg of antibody to H3K27ac (ab4729, Abcam) in the SX-8G IP-Star Automated System (Diagenode) in RIPA buffer (TE with 0.1% SDS, 1% Triton X-100, 150 mM NaCl and 0.1% sodium deoxycholate) following the protocol of the manufacturer. Fastq files from the Illumina pipeline were aligned by Bowtie (version 1.0.0)⁶¹, and unique reads were identified with no more than one error allowed for alignment. MACS2 (version 2.0.10.20130712)⁶² was used to identify islands

of enrichment (q -value threshold less than 0.2) and estimated fragment size. For visualization, data were uploaded to the UCSC Genome Browser.

Calpain activity assays. EPC2 cultures were left untreated or stimulated with IL-13 (100 ng/ml) for 48 h and lysed with Mammalian Protein Extraction Reagent lysis buffer (Pierce, 78501) in the presence of 1 mM EDTA and 1 mM dithiothreitol on ice. Calpain activity was measured with the Promega Calpain-Glo protease assay according to the manufacturer's instructions. Briefly, 50 μl of lysate was incubated for 10 min with reaction buffer with or without 100 mM PD150606 (Santa Cruz Biotechnology, sc-222133), and luminescence was read using the Biotek Synergy2 Multi-Mode microplate reader.

54. Prahalad, S. *et al.* Juvenile rheumatoid arthritis: linkage to HLA demonstrated by allele sharing in affected sibpairs. *Arthritis Rheum.* **43**, 2335–2338 (2000).
55. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
56. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
57. Zimmermann, N. *et al.* Dissection of experimental asthma with DNA microarray analysis identifies arginase in asthma pathogenesis. *J. Clin. Invest.* **111**, 1863–1874 (2003).
58. Flicek, P. *et al.* Ensembl 2013. *Nucleic Acids Res.* **41**, D48–D55 (2013).
59. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
60. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
61. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
62. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).